

*Dedicated to the Memory of Moshe Shaked*

# A CENTRAL LIMIT THEOREM FOR COSTS IN BULINSKAYA'S INVENTORY MANAGEMENT PROBLEM WHEN DELIVERIES FACE DELAYS

ALESSANDRO ARLOTTO AND J. MICHAEL STEELE

**ABSTRACT.** It is common in inventory theory to consider policies that minimize the expected cost of ordering and holding goods or materials. Nevertheless, the realized cost is a random variable, and, as the Saint Petersburg Paradox reminds us, the expected value does not always capture the full economic reality of a decision problem. Here we take the classic inventory model of Bulinskaya (1964), and, by proving an appropriate central limit theorem, we show in a reasonably rich (and practical) sense that the mean-optimal policies are economically appropriate. The motivation and the tools are applicable to a large class of Markov decision problems.

**MATHEMATICS SUBJECT CLASSIFICATION (2010):** Primary: 60C05, 90B05; Secondary: 60F05, 60J05, 90C39, 90C40.

**KEY WORDS:** inventory management, Markov decision problems, central limit theorem, non-homogeneous Markov chain, Dobrushin coefficient, stochastic order, discrete-time martingale.

## 1. WEIGHING COSTS WITH MORE THAN MEANS

In many problems of operations management, the quantity of central interest is the cost  $\mathcal{C}_n(\pi_n^*)$  that one incurs over  $n$  periods by following a policy  $\pi_n^*$  that is optimal in the sense that

$$(1) \quad \mathbb{E}[\mathcal{C}_n(\pi_n^*)] = \min \{ \mathbb{E}[\mathcal{C}_n(\pi)] : \pi \in \mathcal{S}_n \},$$

where  $\mathcal{S}_n$  is a specified set of feasible policies. This procedure is so commonplace that one may see nothing odd in it. Nevertheless, just the mention of the Saint Petersburg Paradox is enough to remind one that economic decisions based on means alone may be quite ill founded. The realized cost  $\mathcal{C}_n(\pi)$  incurred by a policy  $\pi$  is a random variable, and, to compare this to the cost  $\mathcal{C}_n(\pi')$  incurred by another policy  $\pi'$ , one must make an economically meaningful comparison between the distributions of two random variables.

Perhaps no individual has contributed more to our understanding of stochastic orders than Moshe Shaked. The theory of orders on distributions was a life-long interest of his, and this interest also encompassed applications to many fields. Among his contribution that single out applications in economics, one can mention at least Müller et al. (2002), Colangelo et al. (2005), Li and Shaked (2007) and Shaked et al. (2012). The two further papers of Marshall and Shaked (1983)

---

A. Arlotto: The Fuqua School of Business, Duke University, 100 Fuqua Drive, Durham, NC, 27708. Email address: [alessandro.arlotto@duke.edu](mailto:alessandro.arlotto@duke.edu) .

J. M. Steele: Department of Statistics, The Wharton School, University of Pennsylvania, 3730 Walnut Street, Philadelphia, PA, 19104. Email address: [steele@wharton.upenn.edu](mailto:steele@wharton.upenn.edu) .

and Hartman et al. (2000) make specific note of applications to inventory management which is our principal focus here. More broadly, the classic texts of Shaked and Shanthikumar (1994, 2007) have served for a generation as fundamental resources for information on stochastic orders — in their many varieties and for their many applications.

Here, we focus almost entirely on a well-studied inventory management model that was introduced by Bulinskaya (1964), and the first observation is that it leads one into the same kind of economic quandary that one finds with the generic mean-focused optimality criterion (1). In Section 6, we will discuss how one might approach this quandary using the theory of stochastic order, but, in the main, we stick to an approach that is decidedly more classical.

As a first cut, one should note that the criterion (1) becomes economically meaningful precisely in those circumstances where  $\mathbb{E}[\mathcal{C}_n(\pi_n^*)]$  is a viable surrogate for  $\mathcal{C}_n(\pi_n^*)$ , and, at a bare minimum, this calls for  $\mathbb{E}[\mathcal{C}_n(\pi_n^*)]$  and  $\mathcal{C}_n(\pi_n^*)$  to be probabilistically of the same order. This condition does not fully escape the Saint Petersburg Paradox, but it does give one some practical guidance. Moreover, it was found in Arlotto et al. (2014) that such closeness in probability holds automatically for a large class of Markov decision problems.

To take this logic one step further, one can look to determine the asymptotic distribution of the realized cost  $\mathcal{C}_n(\pi_n^*)$ . This is the approach taken here. Specifically, we prove that there is a central limit theorem (CLT) for  $\mathcal{C}_n(\pi_n^*)$  in the Bulinskaya inventory model where one may suffer delivery delays according to chance. Unfortunately, such a CLT still does not free us entirely from the logical tensions of the Saint Petersburg Paradox. Nevertheless, it brings us one step closer to a compelling practical rationale for embracing the mean-optimal policy in Bulinskaya’s famous model.

### 1.1. BULINSKAYA’S INVENTORY MODEL: DELIVERY DELAYS INCLUDED

In the classic dynamic inventory model of Bulinskaya (1964), one has  $n$  periods and  $n$  demands  $D_1, D_2, \dots, D_n$  that are modeled as independent random variables with a common density  $\psi$  that has support on a bounded subinterval of  $[0, \infty)$ . At the beginning of each period  $i \in [1 : n] \equiv \{1, 2, \dots, n\}$  one knows the current level of inventory  $x$ , and the main task is to decide upon the quantity that should be ordered. Here it is technically useful to allow  $x$  to take both positive and negative values, and in the latter case, the absolute value  $|x|$  is understood to represent the level of *backlogged* demand. To underscore this possibility, we sometimes call  $x$  the *generalized* inventory level.

Bulinskaya’s model also allows for some uncertainty about the time when an order is filled. Specifically, the model posits that there is a fixed, known parameter  $q \in [0, 1]$  such that after an order is placed it is filled immediately with probability  $q$ , but, if it is not filled immediately, then it is guaranteed to be filled at the beginning of the next period. The associated fulfillment events are assumed to be independent, and we account for them with independent Bernoulli random variables  $Y_1, Y_2, \dots, Y_n$  where the event  $Y_i = 1$  indicates that the  $i$ ’th order is filled immediately.

Bulinskaya’s model further accounts for costs of two kinds. First, there is the *cost of ordering*, and the model takes this cost to be proportional to the size of the order. Specifically, to move the inventory from level  $x$  to the level  $y \geq x$ , one places

an order of size  $y - x$ , and one incurs a cost of  $c(y - x)$  where the multiplicative constant  $c > 0$  is a model parameter that we assume to be known.

Second, there is the *carrying cost* that one incurs either by holding some physical inventory or by managing a backlog. To model this cost, one introduces two non-negative constants  $c_h$  and  $c_p$ , and one defines a piecewise linear function  $L : \mathbb{R} \rightarrow \mathbb{R}$  by setting

$$(2) \quad L(z) = \begin{cases} c_h z & \text{if } z \geq 0 \\ -c_p z & \text{if } z < 0. \end{cases}$$

The interpretation here is that if  $z \geq 0$ , then  $L(z)$  represents the cost for holding a quantity  $z$  of inventory from one period to the next, while, if  $z < 0$ , then  $L(z)$  represents the penalty cost for managing a backlog of size  $-z = |z| \geq 0$ .

Given this framework, the carrying cost that is realized in period  $i$  depends on whether the order is filled immediately or not. If the inventory at the beginning of period  $i$  is  $x$  and one places an order to bring the inventory up to the level  $y$ , then in the case (i) in which the order is filled immediately the carrying cost for period  $i$  is  $L(y - D_i)$ , but in the case (ii) in which the delivery is delayed to the beginning of the next period the carrying cost is  $L(x - D_i)$ . Thus, the carrying cost for period  $i$  can always be written as

$$(3) \quad Y_i L(y - D_i) + (1 - Y_i) L(x - D_i).$$

Finally, we assume that the ordering cost rate  $c$  is strictly smaller than the backlog management penalty rate  $c_p$ ; thus, it is never optimal to accrue backlog penalty costs if one has the opportunity to place an order.

## 1.2. A MEAN-OPTIMAL POLICY

Bulinskaya studied this problem in great detail, and she characterized the structure of the ordering policy that minimizes the total expected inventory costs over the  $n$  periods. That is, she found the optimal policy in the sense of (1).

In brief, she found values  $\gamma_{n,i}(x)$  such that if the inventory at the beginning of period  $i$  is  $x$ , then it is optimal in the sense of (1) to place an order of size  $\gamma_{n,i}(x) - x$  so that the order would bring the inventory up to the level  $\gamma_{n,i}(x)$ . We call the map  $x \mapsto \gamma_{n,i}(x)$  the *order-up-to function*, and Bulinskaya gave a concrete formula for its value.

Specifically, she found that if one sets

$$(4) \quad n_0 = \inf\{j \in \mathbb{N} : c < c_p(q + j + 1)\}$$

then there is a sequence of non-decreasing real values

$$(5) \quad 0 < s_{n_0+2} \leq s_{n_0+3} \leq \dots \leq s_n \leq s_\infty$$

such for  $i$  in the *principal range*  $[1 : n - n_0] \equiv \{1, 2, \dots, n - n_0 - 1\}$  one has the *active rule*

$$(6) \quad \gamma_{n,i}(x) = \begin{cases} s_{n-i+1} & \text{if } x \leq s_{n-i+1} \\ x & \text{if } x > s_{n-i+1}. \end{cases}$$

Instead, for  $i$  in the *residual range*  $[n - n_0 : n]$ , one has the *passive rule*

$$(7) \quad \gamma_{n,i}(x) = x,$$

so, for  $i$  in the residual range the optimal action is to passively order nothing.

In general, the concrete values of the levels  $\{s_i : n_0 + 2 \leq i \leq n\}$  can only be computed numerically, but those numerical values are not needed here. Parenthetically, we should also note that the curious starting level  $s_{n_0+2}$  comes about because it is the value that one needs to compute to characterize  $\gamma_{n,i}(x)$  when  $i = n - n_0 - 1$ ; i.e. when  $i$  is the largest element of the principal range  $[1 : n - n_0]$ .

Bulinskaya's optimal restocking rules determine a natural *inventory process*, where we view  $X_{n,i}$  as the generalized inventory at the *beginning of period  $i$* , which we define to be the moment before one places the order for period  $i$ . If we take the initial inventory level to be  $X_{n,1} = x$ , then the restocking rules (6) and (7) tell us that the subsequent value of the generalized inventory level at the beginning of period  $i + 1$  is always given by

$$(8) \quad X_{n,i+1} = \gamma_{n,i}(X_{n,i}) - D_i \quad \text{for all } i \in [1 : n].$$

In particular, the generalized inventory at the beginning of period  $i + 1$  does not depend on whether the order in period  $i$  was delayed or not. In fact, a period  $i$  order that is delayed is assumed to be filled by the beginning of period  $i + 1$ .

From the recursion (8), we see that the sequence  $\{X_{n,i} : 1 \leq i \leq n + 1\}$  is a non-homogenous Markov chain, and, if we let  $\mathcal{F}_0$  denote the trivial  $\sigma$ -field and set

$$(9) \quad \mathcal{F}_i = \sigma\{D_1, D_2, \dots, D_i, Y_1, Y_2, \dots, Y_i\}, \quad \text{for } i \in [1 : n],$$

then we see that  $X_{n,i}$  is  $\mathcal{F}_{i-1}$ -measurable for each  $i \in [1 : n + 1]$ .

Now, if we take  $0 < J < \infty$  to be the least upper bound of the support of the demand distribution  $\psi$  and if we begin with  $X_{n,1} = x \in [0, s_\infty]$ , then the recursion (8) tells us that

$$\begin{aligned} X_{n,i} &\in [-J, s_\infty] \quad \text{for all } i \in [2 : n - n_0] \text{ and} \\ X_{n,i} &\in [-(n_0 + 2)J, s_\infty] \quad \text{for all } i \in [n - n_0 : n + 1]. \end{aligned}$$

The most natural starting inventory level is  $X_{n,1} = 0$ , and, *de minimus*, we will assume throughout that  $X_{n,1} = x \in [0, s_\infty]$ . As a consequence, we can take the state space of the generalized inventory process to be  $\mathcal{X} = [-(n_0 + 2)J, s_\infty]$ . So, in particular, we have that the state space  $\mathcal{X}$  of the inventory chain is bounded.

### 1.3. REALIZED COSTS

The restocking policy  $\pi_n^*$  determined by Bulinskaya's rules (5) and (6) is optimal in the sense that it attains the minimal *expected* inventory cost over all feasible policies. Of course, the actual cost realized by following the policy  $\pi_n^*$  is a random variable, and it turns out to be useful to write it as the cumulative sum of the ordering cost and the carrying costs.

The *one-period cost*  $\mathcal{P}_{n,i}$  for period  $i$  is the sum of (i) the ordering costs and (ii) the holding cost that one calculates from (3). Thus, when we follow Bulinskaya's mean-optimal inventory policy, the one-period cost for period  $i$  is given by

$$(10) \quad \mathcal{P}_{n,i} = c(\gamma_{n,i}(X_{n,i}) - X_{n,i}) + Y_i L(\gamma_{n,i}(X_{n,i}) - D_i) + (1 - Y_i)L(X_{n,i} - D_i).$$

Thus, the corresponding *realized cost* accrued over time is

$$(11) \quad \mathcal{C}_i(\pi_n^*) = \sum_{j=1}^i \mathcal{P}_{n,j}, \quad \text{for } i \in [1 : n],$$

and, of course, our main concern is with the total cost  $\mathcal{C}_n(\pi_n^*)$  over the full time horizon.

As we noted before, the mean-focused optimality (1) of Bulinskaya's policy is economically meaningful only if the expected value of the total cost  $\mathcal{C}_n(\pi_n^*)$  can serve as a good proxy for the cost that is actually realized. While there is no *a priori* reason for this to be the case, we can increase our confidence in the use of  $\pi_n^*$  if we learn more about the distribution of the total realized cost  $\mathcal{C}_n(\pi_n^*)$ .

#### 1.4. A CLT MEAN-OPTIMAL FOR INVENTORY COSTS

For Bulinskaya's optimal policy one can show that after the natural centering and scaling the total realized cost  $\mathcal{C}_n(\pi_n^*)$  is asymptotically normal — provided that the demand density satisfies a natural regularity condition.

To be explicit, we say that a demand density  $\psi$  is *softly unimodal* provided that for each  $\epsilon \geq 0$  there is a  $\hat{w} = \hat{w}(\epsilon) \in [-\infty, \infty)$  such that

$$(12) \quad \{w : \psi(w + \epsilon) \leq \psi(w)\} = [\hat{w}, \infty).$$

Naturally any unimodal density is softly unimodal, and any density with two isolated modes is not softly unimodal. The uniform density on  $[0, 1]$  is a simple example of a density which is not unimodal but which is softly unimodal. Although the condition (12) is a little unconventional, it is satisfied by essentially all of the distributions that are commonly considered in inventory theory. Moreover, it is exactly what is needed in our context.

**Theorem 1** (CLT for Mean-Optimal Inventory Cost). *If the demand density  $\psi$  is softly unimodal and has bounded support in  $[0, \infty)$ , then the inventory cost  $\mathcal{C}_n(\pi_n^*)$  that is realized under the mean-optimal policy  $\pi_n^*$  satisfies*

$$\frac{\mathcal{C}_n(\pi_n^*) - \mathbb{E}[\mathcal{C}_n(\pi_n^*)]}{\sqrt{\text{Var}[\mathcal{C}_n(\pi_n^*)]}} \implies N(0, 1), \quad \text{as } n \rightarrow \infty.$$

We will prove this theorem over the next several sections. In the first of these, we recall a CLT for non-homogenous Markov chains from Arlotto and Steele (2016), where there is also a discussion of the CLT for  $\mathcal{C}_n(\pi_n^*)$  in the context of the much simpler Bulinskaya model where one does not have delayed deliveries.

Our main arguments are distributed over Sections 3, 4, and 5. Finally, in Section 6 we note three further steps that one can take toward the economic justification for the use of the mean-optimal policies in Bulinskaya's model and related problems. One of these depends on the theory of stochastic orders.

## 2. BACKGROUND ON NON-HOMOGENEOUS MARKOV CHAINS

The proof of Theorem 1 depends on a recent theorem for functions on the path of a finite horizon, temporally non-homogenous Markov chain. Here, we generically denote such a chain by  $\{Z_{N,i} : 1 \leq i \leq N\}$ . We also assume that the horizon  $N \in \mathbb{N}$  is fixed, and the state space  $\mathcal{Z}$  of the chain is a Borel space.

If  $\mathcal{B}(\mathcal{Z})$  denotes the class of Borel subsets of  $\mathcal{Z}$  and  $\nu$  is the initial distribution of the chain, then we have the relation  $\nu(B) = \mathbb{P}(Z_{N,1} \in B)$  for all  $B \in \mathcal{B}(\mathcal{Z})$ . The laws for the subsequent values of the chain are then determined by specifying the one-step transition kernels  $\{K_{i,i+1}^{(N)} : 1 \leq i < N\}$  such that

$$K_{i,i+1}^{(N)}(z, B) = \mathbb{P}(Z_{N,i+1} \in B \mid Z_{N,i} = z) \quad \text{for } z \in \mathcal{Z} \text{ and } B \in \mathcal{B}(\mathcal{Z}).$$

While the transition kernels can be quite general, we do require some conditions. To state these we first recall that for any Markov transition kernel  $K$  on  $\mathcal{Z}$ , the

*Dobrushin contraction coefficient* of  $K$  is defined by

$$(13) \quad \delta(K) = \sup_{\substack{z_1, z_2 \in \mathcal{Z} \\ B \in \mathcal{B}(\mathcal{Z})}} |K(z_1, B) - K(z_2, B)|,$$

and the corresponding *ergodic coefficient* of  $K$  is given by

$$\alpha(K) = 1 - \delta(K).$$

Our main concern here is with functions on the paths of an array of time non-homogeneous Markov chains. To define these functions we first fix an integer  $m$  and for each  $n = 1, 2, \dots$  we introduce an array of  $n$  real valued functions of  $1 + m$  variables,

$$\varphi_{n,i} : \mathcal{Z}^{1+m} \rightarrow \mathbb{R}, \quad i \in [1 : n].$$

If we now set  $N = n + m$ , then the object of central interest is the sum

$$(14) \quad S_n = \sum_{i=1}^n \varphi_{n,i}(Z_{N,i}, \dots, Z_{N,i+m}).$$

Here, one should note that this sum with  $n$  summands is a function of the  $N$  variables  $\{Z_{N,i} : 1 \leq i \leq N\}$  where  $N = n + m$ .

## 2.1. A GENERAL CENTRAL LIMIT THEOREM

In favorable circumstances, when the sums  $\{S_n : n \geq 1\}$  defined by (14) are centered about their mean and scaled by their standard deviations, one would naturally expect that they converge in distribution to the standard Gaussian. Arlotto and Steele (2016) proved that this is indeed the case provided that one has a modest compatibility relation between the sizes of three quantities: (i) the sizes of the ergodic coefficients  $\alpha(K_{i,i+1}^{(N)})$ , (ii) the functions  $\{\varphi_{n,i} : 1 \leq i \leq n\}$ , and (iii) the variance of  $S_n$ . Dobrushin (1956) and Sethuraman and Varadhan (2005) had earlier proved theorems that cover the special case where  $m = 0$  (or  $N = n$ ).

To write that compatibility relation compactly for general arrays of transition kernels, it is useful to introduce one further quantity. For given values of  $m$ ,  $n$ , and  $N = n + m$ , the key quantity in the compatibility relation is given by

$$\alpha_n = \min_{1 \leq i < n} \alpha(K_{i,i+1}^{(N)}).$$

There should be no cause for confusion if we call this the  $n$ 'th *minimal ergodic coefficient*, since, in the special case with  $m = 0$ , it reduces to the classical quantity of the same name. The benefit of  $\alpha_n$  is that gives us an exceptionally useful measure of the “randomness” of the process determined by the  $N$ 'th row of the array of transition kernels. Here and subsequently we use  $\|f\|_\infty$  to denote the essential supremum of the absolute value of the function  $f$ .

**Theorem 2** (Arlotto and Steele, 2016). *If there are constants  $C_1, C_2, \dots$  such that*

$$\max_{1 \leq i \leq n} \|\varphi_{n,i}\|_\infty \leq C_n \quad \text{and} \quad C_n^2 \alpha_n^{-2} = o(\text{Var}[S_n]),$$

*then one has the convergence in distribution*

$$(15) \quad \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\text{Var}[S_n]}} \Rightarrow N(0, 1), \quad \text{as } n \rightarrow \infty.$$

**Corollary 3.** *If there are constants  $c > 0$  and  $C < \infty$  such that*

$$\alpha_n \geq c \quad \text{and} \quad C_n \leq C \quad \text{for all } n \geq 1,$$

*then one has the asymptotic normality (15) whenever  $\text{Var}[S_n] \rightarrow \infty$  as  $n \rightarrow \infty$ .*

### 3. CONNECTING THEOREMS 1 AND 2

The key observation is that with some restructuring the realized inventory cost  $\mathcal{C}_n(\pi_n^*)$  in (11) can be seen as a sum of the form (14). In particular, we need to construct an appropriate temporally non-homogeneous Markov chain.

We do not apply Theorem 2 to the inventory process (8), but rather to a related process on the enlarged state space  $\mathcal{Z} \stackrel{\text{def}}{=} \mathcal{X} \times \{0, 1\}$  that we define by setting

$$Z_{N,i} = (X_{n,i}, Y_i) \quad \text{for } 1 \leq i \leq N = n + 1.$$

If we now set

$$\varphi_{n,i}((x, y), (x', y')) = c(\gamma_{n,i}(x) - x) + yL(x') + (1 - y)L(x - \gamma_{n,i}(x) + x'),$$

then by (10) the one period cost equals  $\varphi_{n,i}(Z_{N,i}, Z_{N,i+1})$  and the total cost  $\mathcal{C}_n(\pi_n^*)$  defined by (11) has the representation

$$\mathcal{C}_n(\pi_n^*) = \sum_{i=1}^n \varphi_{n,i}(Z_{N,i}, Z_{N,i+1}).$$

We noted in Subsection 1.2 that the state space  $\mathcal{X}$  is bounded, so by the piecewise linearity of  $L$ , we see that the cost functions  $\{\varphi_{n,i} : 1 \leq i \leq n\}$  are uniformly bounded.

Now, recalling the principal range  $[1 : n - n_0]$  and the residual range  $[n - n_0 : n]$  that determine the restocking rules (6) and (7), we consider the decomposition

$$(16) \quad \mathcal{C}_n(\pi_n^*) = \sum_{i=1}^{n-n_0-1} \varphi_{n,i}(Z_{N,i}, Z_{N,i+1}) + \sum_{i=n-n_0}^n \varphi_{n,i}(Z_{N,i}, Z_{N,i+1}) \\ \stackrel{\text{def}}{=} \mathcal{C}'_n(\pi_n^*) + \mathcal{C}''_n(\pi_n^*).$$

The sum  $\mathcal{C}''_n(\pi_n^*)$  has  $n_0 + 1$  terms, and we know by (4) that  $n_0$  is bounded by a constant that does not depend on  $n$ . Since the summands of  $\mathcal{C}''_n(\pi_n^*)$  are uniformly bounded, we see that the norm  $\|\mathcal{C}''_n(\pi_n^*)\|_\infty$  is also bounded independently of  $n$ . Consequently, if  $\text{Var}[\mathcal{C}'_n(\pi_n^*)] \rightarrow \infty$  and if  $\mathcal{C}'_n(\pi_n^*)$  satisfies the CLT, then we will find that  $\mathcal{C}_n(\pi_n^*)$  also satisfies the same CLT.

To prove the CLT for  $\mathcal{C}'_n(\pi_n^*)$  there are two basic steps — neither of which is immediate. First, one needs an appropriate lower bound on the minimal ergodic coefficient of the chain  $\{Z_{N,i} : 1 \leq i \leq n - n_0\}$ . Second, one needs to show that the variance of the first sum  $\mathcal{C}'_n(\pi_n^*)$  grows to infinity as  $n \rightarrow \infty$ .

### 4. STEP ONE: LOWER BOUNDS FOR ERGODIC COEFFICIENTS

Given the construction of the inventory process  $\{X_{n,1}, X_{n,2}, \dots, X_{n,n+1}\}$ , we can define one-step transition kernels on the state space  $\mathcal{X}$  by setting

$$K_{i,i+1}^{(n)}(x, A) = \mathbb{P}(X_{n,i+1} \in A \mid X_{n,i} = x)$$

for all  $x \in \mathcal{X}$ ,  $A \in \mathcal{B}(\mathcal{X})$  and  $i \in [1 : n]$ . Similarly, for our newly constructed process  $\{Z_{N,i} : 1 \leq i \leq N\}$  we have one-step transition kernels on  $\mathcal{Z} = \mathcal{X} \times \{0, 1\}$  that we can write as

$$\widehat{K}_{i,i+1}^{(N)}(z, B) = \mathbb{P}(Z_{N,i+1} \in B \mid Z_{N,i} = z),$$

where  $z \in \mathcal{Z}$ ,  $B \in \mathcal{B}(\mathcal{Z})$ , and  $i \in [1 : N]$ .

If we set  $B_0 = B \cap \{z = (x, y) : y = 0\}$  and  $B_1 = B \cap \{z = (x, y) : y = 1\}$  then  $B_0$  and  $B_1$  are disjoint Borel subsets of  $\mathcal{Z}$  with union  $B$ . Thus, if we also set  $A_0 = \{x : (x, 0) \in B_0\}$  and  $A_1 = \{x : (x, 1) \in B_0\}$ , then by the independence assumptions from Section 3 we have

$$\begin{aligned} (17) \quad \widehat{K}_{i,i+1}^{(N)}(z, B) &= \widehat{K}_{i,i+1}^{(N)}(z, B_0) + \widehat{K}_{i,i+1}^{(N)}(z, B_1) \\ &= (1 - q)K_{i,i+1}^{(n)}(x, A_0) + qK_{i,i+1}^{(n)}(x, A_1). \end{aligned}$$

By the decomposition (17) and the definition (13) of the Dobrushin coefficient, we now have a remarkably nice relationship:

$$(18) \quad \delta(\widehat{K}_{i,i+1}^{(N)}) \leq \delta(K_{i,i+1}^{(n)}).$$

From (18) we see that we can direct our efforts toward obtaining a uniform upper bound on  $\delta(K_{i,i+1}^{(n)})$ . We begin this analysis with general lemma that also explains how the soft unimodality condition (12) helps us.

**Lemma 4.** *If the density  $\psi$  of  $D_1$  satisfies the soft unimodality condition (12), then for  $\epsilon_i = |\gamma_{n,i}(x') - \gamma_{n,i}(x)|$  one has*

$$(19) \quad \sup_{A \in \mathcal{B}(\mathcal{X})} |K_{i,i+1}^{(n)}(x', A) - K_{i,i+1}^{(n)}(x, A)| = \mathbb{P}(\widehat{w} \leq D_1 \leq \widehat{w} + \epsilon_i),$$

where  $\widehat{w} = \widehat{w}(\epsilon_i)$  is the value guaranteed by the condition (12).

*Proof.* Given  $x \in \mathcal{X}$  and a Borel set  $A \subseteq \mathcal{X}$ , we introduce the Borel set

$$A_x = \gamma_{n,i}(x) - A,$$

so that the transition kernel  $K_{i,i+1}^{(n)}(x, A)$  can be written as

$$K_{i,i+1}^{(n)}(x, A) = \mathbb{P}(X_{n,i+1} \in A \mid X_{n,i} = x) = \mathbb{P}(D_1 \in A_x) = \int_{A_x} \psi(w) dw.$$

Without loss of generality we can assume that  $x \leq x'$ , so the restocking formula (6) gives us  $\gamma_{n,i}(x) \leq \gamma_{n,i}(x')$ , and for  $\epsilon_i = \gamma_{n,i}(x') - \gamma_{n,i}(x) \geq 0$  we find

$$K_{i,i+1}^{(n)}(x', A) = \mathbb{P}(X_{n,i+1} \in A \mid X_{n,i} = x') = \mathbb{P}(D_1 - \epsilon_i \in A_x) = \int_{A_x} \psi(w + \epsilon_i) dw.$$

Taking the difference of the last two displays gives us

$$\begin{aligned} |K_{i,i+1}^{(n)}(x', A) - K_{i,i+1}^{(n)}(x, A)| &= \left| \int_{A_x} \psi(w) dw - \int_{A_x} \psi(w + \epsilon_i) dw \right| \\ &\leq \sup_{C \in \mathcal{B}(\mathbb{R})} \left| \int_C \psi(w) dw - \int_C \psi(w + \epsilon_i) dw \right|. \end{aligned}$$



Here the supremum is attained at  $C^* = \{w : \psi(w + \epsilon_i) \leq \psi(w)\}$ , and by the soft unimodality condition (12) for  $\psi$ , we have  $C^* = [\hat{w}, \infty)$ . Hence, we have

$$\begin{aligned} \sup_{A \in \mathcal{B}(\mathcal{X})} |K_{i,i+1}^{(n)}(x', A) - K_{i,i+1}^{(n)}(x, A)| &\leq \int_{\hat{w}}^{\infty} \{\psi(x) - \psi(x + \epsilon_i)\} dx \\ &= \mathbb{P}(D_1 \geq \hat{w}) - \mathbb{P}(D_1 - \epsilon_i \geq \hat{w}) = \mathbb{P}(\hat{w} \leq D_1 \leq \hat{w} + \epsilon_i), \end{aligned}$$

just as needed to complete the proof of the lemma.  $\square$

Now we need a uniform estimate for the bounding probability in (19). An explicit bound can be given only over the principal range, and this is the reason why we work with the decomposition (16). The argument is not difficult but it requires some further facts from Bulinskaya (1964).

**Lemma 5.** *If  $x, x' \in \mathcal{X}$ ,  $\epsilon_i = |\gamma_{n,i}(x') - \gamma_{n,i}(x)|$ , and*

$$\kappa = \max \left\{ \frac{c_p}{c_h + c_p}, \frac{(q + n_0 + 1)c_h + c}{(q + n_0 + 1)(c_h + c_p)} \right\},$$

*then for all  $i \in [1 : n - n_0)$  one has*

$$\sup_{w \in \mathbb{R}} \mathbb{P}(w \leq D_1 \leq w + \epsilon_i) \leq \kappa < 1.$$

*Proof.* Without any loss of generality, we again take  $x \leq x'$ . Next, by the monotonicity of the restocking formula (6) and the monotonicity of the optimal stocking levels (5) we have the bounds

$$(20) \quad s_{n_0+2} \leq \gamma_{n,i}(x) \leq \gamma_{n,i}(x') \leq s_{\infty} \quad \text{for all } i \in [1 : n - n_0).$$

For  $\epsilon_i = \gamma_{n,i}(x') - \gamma_{n,i}(x)$  this tells us

$$0 \leq \epsilon_i = \gamma_{n,i}(x') - \gamma_{n,i}(x) \leq s_{\infty} - s_{n_0+2}.$$

If  $w + \epsilon_i \leq s_{\infty}$ , then we have the trivial bound

$$(21) \quad \mathbb{P}(w \leq D_i \leq w + \epsilon_i) \leq \mathbb{P}(D_i \leq s_{\infty}).$$

On the other hand, if  $w + \epsilon_i \geq s_{\infty}$  then by (20) we get  $w \geq s_{n_0+2}$  and we again have a trivial bound,

$$(22) \quad \mathbb{P}(w \leq D_i \leq w + \epsilon_i) \leq \mathbb{P}(D_i \geq s_{n_0+2}).$$

To make use of these two bounds, we need some estimates on the stocking levels  $s_{n_0+2}$  and  $s_{\infty}$ . Fortunately, from Bulinskaya (1964, Theorems 1 and 2) we find that for  $\Psi(t) = \mathbb{P}(D_1 \leq t)$  one has

$$s_{n_0+2} \geq \Psi^{-1} \left( \frac{(q + n_0 + 1)c_p - c}{(q + n_0 + 1)(c_p + c_h)} \right) \quad \text{and} \quad s_{\infty} \leq \Psi^{-1} \left( \frac{c_p}{c_p + c_h} \right),$$

or, in other words,

$$(23) \quad \mathbb{P}(D_i \leq s_{\infty}) \leq \frac{c_p}{c_h + c_p} \quad \text{and} \quad \mathbb{P}(D_i \geq s_{n_0+2}) \leq \frac{(q + n_0 + 1)c_h + c}{(q + n_0 + 1)(c_p + c_h)},$$

Since  $c_h > 0$  the first ratio is always strictly less than one. Also, by the definition (4) of  $n_0$  we have the strict inequality  $c < (q + n_0 + 1)c_p$ , so the second ratio in (23) is strictly less than one. This gives us  $\kappa < 1$ , and, by the bounds (21), (22), and (23) the proof of the lemma is complete.  $\square$

By Lemmas 4 and 5, we see that for each index  $i$  in the principal range  $[1 : n - n_0]$  we have a uniform bound on the Dobrushin coefficient:

$$\delta(K_{i,i+1}^{(n)}) = \sup_{\substack{x, x' \in \mathcal{X} \\ A \in \mathcal{B}(\mathcal{X})}} |K_{i,i+1}^{(n)}(x, A) - K_{i,i+1}^{(n)}(x', A)| \leq \kappa.$$

In turn, the elementary bound (18), tells us in turn that for the minimal ergodic coefficient  $\alpha_{n-n_0}$  of the non-homogeneous Markov chain  $\{Z_{N,i} : 1 \leq i \leq n - n_0\}$  we have

$$\alpha_{n-n_0} \geq \min_{1 \leq i \leq n-n_0} \{1 - \delta(K_{i,i+1}^{(n)})\} \geq 1 - \kappa > 0,$$

and this completes our first step in the proof of the CLT for  $\mathcal{C}'_n(\pi_n^*)$ .

## 5. STEP TWO: VARIANCE LOWER BOUND

Here the plan is to estimate the variance of  $\mathcal{C}_n(\pi_n^*)$  with help from a martingale that we derive from the Bellman equation of dynamic programming. In our case, this is a recursive equation for the values  $\{v_k(x) : x \in \mathcal{X} \text{ and } 1 \leq k \leq n\}$  that represent the expected total inventory cost over the  $k$  time periods in  $[n - k + 1 : n]$  when (i) we begin with an inventory level  $x$  at the beginning of period  $n - k + 1$  and (ii) we follow the optimal policy  $\pi_n^*$  over the remaining  $k$  periods. In dynamic programming the map  $x \mapsto v_k(x)$  is called the value function, or, the *cost-to-go function* since it represents the expected cost under the optimal policy when there are  $k$  remaining periods and the generalized inventory at the beginning of period  $n - k + 1$  is equal to  $x$ .

If we now take  $v_0(x) \equiv 0$  for all  $x \in \mathcal{X}$  as our terminal condition, then we find by backward induction (or, equivalently, by the optimality principle of dynamic programming) that for  $k \in [1 : n]$  the function  $v_k(x)$  can be computed by the recursion:

$$(24) \quad v_k(x) = \min_{y \geq x} \left\{ c(y - x) + \mathbb{E}[Y_{n-k+1}L(y - D_{n-k+1})] \right. \\ \left. + \mathbb{E}[(1 - Y_{n-k+1})L(x - D_{n-k+1})] + \mathbb{E}[v_{k-1}(y - D_{n-k+1})] \right\}.$$

In particular, if the initial inventory level is  $x$ , then one has

$$v_n(x) = \mathbb{E}[\mathcal{C}_n(\pi_n^*)].$$

Bulinskaya (1964) established that the minimum in (24) is uniquely attained at the value  $y = \gamma_{n,n-k+1}(x)$ . This is the order-up-to level for period  $n - k + 1$ , so the inventory at the beginning of period  $n - k + 2$  is given by  $\gamma_{n,n-k+1}(x) - D_{n-k+1}$ . Thus, from the beginning of that period forward, the expected remaining cost under the optimal strategy  $\pi_n^*$  is given by  $v_{k-1}(\gamma_{n,n-k+1}(x) - D_{n-k+1})$ .

The recursion (24) is usefully viewed as the sum of two pieces; specifically, we write

$$(25) \quad v_k(x) = A_k(x) + B_k(x)$$

where we define  $A_k(\cdot)$  and  $B_k(\cdot)$  by setting

$$A_k(x) = c(\gamma_{n,n-k+1}(x) - x) + \mathbb{E}[(1 - Y_{n-k+1})L(x - D_{n-k+1})], \quad \text{and} \\ B_k(x) = \mathbb{E}[Y_{n-k+1}L(\gamma_{n,n-k+1}(x) - D_{n-k+1})] + \mathbb{E}[v_{k-1}(\gamma_{n,n-k+1}(x) - D_{n-k+1})].$$

By the analysis of Bulinskaya (1964), the order-up-to function  $\gamma_{n,n-k+1}(x)$  is given more explicitly by the restocking rules (6) and (7), but the recursion (25) is independently useful. For example, we can use it now to prove a technical lemma that we will need shortly.

**Lemma 6.** *If  $x \leq 0$ ,  $x' \leq 0$ , and  $n - k + 1 \in [1 : n - n_0]$ , then one has*

$$v_k(x) - v_k(x') = (x' - x)\{c + c_p(1 - q)\}.$$

*Proof.* The crucial observation here is that for  $n - k + 1$  in the principal range  $[1 : n - n_0]$ , the restocking rule (6) tells us that for  $x$  and  $x'$  in  $(-\infty, 0]$  we have

$$(26) \quad \gamma_{n,n-k+1}(x) = \gamma_{n,n-k+1}(x') = s_k.$$

Thus, the formula for  $B_k(\cdot)$  in the decomposition (25) gives us  $B_k(x) = B_k(x')$ . By the definition (2) of the carrying-cost function  $L$ , the independence of  $Y_{n-k+1}$  and  $D_{n-k+1}$ , and a second application of (26), we then have

$$\begin{aligned} v_k(x) - v_k(x') &= A_k(x) - A_k(x') \\ &= c(x' - x) + (1 - q)\mathbb{E}[(-c_p)(x - D_{n-k+1}) + c_p(x' - D_{n-k+1})] \\ &= (x' - x)\{c + c_p(1 - q)\}, \end{aligned}$$

just as needed to complete the proof of the lemma.  $\square$

A more thematic benefit of the value functions  $\{v_k : 0 \leq k \leq n\}$  is that they lead one to a useful martingale with respect to the natural filtration  $\{\mathcal{F}_i : 0 \leq i \leq n\}$  that was introduced in (9).

**Proposition 7** (Optimality Martingale). *The process defined for  $i \in [0 : n]$  by*

$$(27) \quad \mathcal{M}_{n,i} = \mathcal{C}_i(\pi_n^*) + v_{n-i}(X_{n,i+1}) = \mathcal{C}_i(\pi_n^*) + v_{n-i}(\gamma_{n,i}(X_{n,i}) - D_i),$$

*is an  $\{\mathcal{F}_i : 0 \leq i \leq n\}$  adapted martingale. The martingale differences*

$$(28) \quad d_{n,i} \equiv \mathcal{M}_{n,i} - \mathcal{M}_{n,i-1} = \mathcal{P}_{n,i} + v_{n-i}(X_{n,i+1}) - v_{n-i+1}(X_{n,i})$$

*then provide the representation*

$$(29) \quad \mathcal{C}_n(\pi_n^*) - \mathbb{E}[\mathcal{C}_n(\pi_n^*)] = \sum_{i=1}^n d_{n,i}.$$

*Proof.* The  $\mathcal{F}_i$ -measurability of  $\mathcal{M}_{n,i}$  is obvious, and the integrability of  $\mathcal{M}_{n,i}$  follows from the boundedness of state space  $\mathcal{X}$  and the piecewise linearity of  $L$ . Also, the second equality in (27) is simply given by the recursive identity (8).

To get the martingale property, we first note that the definition of the value function gives us  $v_{n-i}(X_{n,i+1}) \equiv \mathbb{E}[\mathcal{C}_n(\pi_n^*) - \mathcal{C}_i(\pi_n^*) | \mathcal{F}_i]$ , so, from the definition of  $\mathcal{M}_{n,i}$ , we have

$$\begin{aligned} \mathcal{M}_{n,i} &= \mathcal{C}_i(\pi_n^*) + v_{n-i}(X_{n,i+1}) = \mathcal{C}_i(\pi_n^*) + \mathbb{E}[\mathcal{C}_n(\pi_n^*) - \mathcal{C}_i(\pi_n^*) | \mathcal{F}_i] \\ &= \mathbb{E}[\mathcal{C}_n(\pi_n^*) | \mathcal{F}_i]. \end{aligned}$$

This confirms that the process  $\{\mathcal{M}_{n,i}, 0 \leq i \leq n\}$  is a martingale with respect to the increasing sequence of  $\sigma$ -fields  $\{\mathcal{F}_i : 0 \leq i \leq n\}$ . It is, in fact, a Doob martingale. Finally, we should note that since  $\mathcal{M}_{n,0} = \mathbb{E}[\mathcal{C}_n(\pi_n^*)]$  and  $\mathcal{M}_{n,n} = \mathcal{C}_n(\pi_n^*)$ , we get the representation (29) from the telescoping sum of the differences.  $\square$

**Remark 8** (Uniform Boundedness Property). One can further prove that there is a constant  $B < \infty$  such that

$$(30) \quad \|d_{n,i}\|_\infty \leq B \quad \text{for all } 1 \leq i \leq n < \infty,$$

and, as we will explain in Section 6, this fact has some useful implications. Nevertheless, we do not pursue the proof of (30) here since it is not needed for the proof of Theorem 1.

We now have all of the tools in place to prove a key lemma. It will lead in turn to the completion of the second step in the proof of the CLT for  $\mathcal{C}'_n(\pi_n^*)$ , and, from that point, the completion of the proof of Theorem 1 is easy.

**Lemma 9.** *There are constants  $\beta > 0$  and  $N < \infty$  such that for all  $n \geq N$ , one has the variance lower bound*

$$(31) \quad \text{Var}[\mathcal{C}_n(\pi_n^*)] = \sum_{i=1}^n \mathbb{E}[d_{n,i}^2] \geq \beta n.$$

*Proof.* Our proof of (31) will depend on bounding the second conditional moment  $\mathbb{E}[d_{n,i}^2 | \mathcal{F}_{i-1}]$  from below, and, since  $X_{n,i}$  is  $\mathcal{F}_{i-1}$ -measurable, we see that when we condition on  $\mathcal{F}_{i-1}$  the random variable  $d_{n,i}$  is a function of just  $Y_i$  and  $D_i$ . If we call this function  $f_{n,i}$ , then we can also write

$$d_{n,i} = f_{n,i}(Y_i, D_i) \quad \text{given } \mathcal{F}_{i-1}.$$

We now consider a pair  $(Y'_i, D'_i)$  such that  $(Y'_i, D'_i)$  is independent of  $(Y_i, D_i)$  and has the same distribution as  $(Y_i, D_i)$ . The random variables  $f_{n,i}(Y_i, D_i)$  and  $f_{n,i}(Y'_i, D'_i)$  are therefore independent. Moreover, conditional on the  $\sigma$ -field  $\mathcal{F}_{i-1}$  they have the same distribution as  $d_{n,i}$  so if we set  $\Delta_i = f_{n,i}(Y_i, D_i) - f_{n,i}(Y'_i, D'_i)$  then by a familiar calculation we have

$$(32) \quad \mathbb{E}[d_{n,i}^2 | \mathcal{F}_{i-1}] = \frac{1}{2} \mathbb{E}[\Delta_i^2 | \mathcal{F}_{i-1}].$$

Using the definition (28) and the inventory recursion (8) can now write

$$(33) \quad f_{n,i}(Y_i, D_i) = c(\gamma_{n,i}(X_{n,i}) - X_{n,i}) + g_{n,i}(Y_i, D_i) + h_{n,i}(Y_i, D_i)$$

where

$$g_{n,i}(Y_i, D_i) = Y_i L(\gamma_{n,i}(X_{n,i}) - D_i) + (1 - Y_i) L(X_{n,i} - D_i)$$

and

$$h_{n,i}(Y_i, D_i) = v_{n-i}(\gamma_{n,i}(X_{n,i}) - D_i) - v_{n-i+1}(X_{n,i}).$$

To work toward a simple lower bound for the conditional expectation (32), we now consider the event

$$G_i = \{ \omega : D_i(\omega) \in [s_\infty, J], D'_i(\omega) \in [s_\infty, J], \text{ and } Y_i(\omega) = Y'_i(\omega) \},$$

where  $J$  is the supremum of the support of the demand density  $\psi$  and  $s_\infty$  is the top stocking level in (5).

Since  $X_{n,i} \leq s_\infty$  and  $\gamma_{n,i}(X_{n,i}) \leq s_\infty$ , on  $G_i$  we have relations

$$(34) \quad X_{n,i+1} = \gamma_{n,i}(X_{n,i}) - D_i \leq 0 \quad \text{and} \quad X'_{n,i+1} = \gamma_{n,i}(X_{n,i}) - D'_i \leq 0.$$

Now, if we use  $\mathbb{1}(A)$  to denote the indicator function of an event  $A$ , then for all  $i \in [1 : n - n_0]$  we have from Lemma 6 that

$$(35) \quad \{h_{n,i}(Y_i, D_i) - h_{n,i}(Y'_i, D'_i)\} \mathbb{1}(G_i) = (D_i - D'_i) \{c + c_p(1 - q)\} \mathbb{1}(G_i).$$

On  $G_i$  we also have  $Y_i = Y'_i$  so by (34) and by the definition (2) of the carrying-cost function  $L$ , we have

$$\begin{aligned}
& \{g_{n,i}(Y_i, D_i) - g_{n,i}(Y'_i, D'_i)\} \mathbb{1}(G_i) \\
&= Y_i \{L(\gamma_{n,i}(X_{n,i}) - D_i) - L(\gamma_{n,i}(X_{n,i}) - D'_i)\} \mathbb{1}(G_i) \\
&\quad + (1 - Y_i) \{L(X_{n,i} - D_i) - L(X_{n,i} - D'_i)\} \mathbb{1}(G_i) \\
&= Y_i \{-c_p(\gamma_{n,i}(X_{n,i}) - D_i) + c_p(\gamma_{n,i}(X_{n,i}) - D'_i)\} \mathbb{1}(G_i) \\
&\quad + (1 - Y_i) \{-c_p(X_{n,i} - D_i) + c_p(X_{n,i} - D'_i)\} \mathbb{1}(G_i) \\
&= c_p(D_i - D'_i) \mathbb{1}(G_i).
\end{aligned}$$

From the last relation, the identity (35), and the definition (33) of  $f_{n,i}$ , we have for all  $i \in [1 : n - n_0]$  that

$$\{f_{n,i}(Y_i, D_i) - f_{n,i}(Y'_i, D'_i)\} \mathbb{1}(G_i) = (D_i - D'_i) \{c + (2 - q)c_p\} \mathbb{1}(G_i).$$

We also have  $\Delta_i = \mathbb{1}(G_i)\Delta_i + \mathbb{1}(G_i^c)\Delta_i$  so if we set  $\rho \equiv \{c + (2 - q)c_p\} > 0$  we have  $\Delta_i^2 \geq \mathbb{1}(G_i)\rho^2(D_i - D'_i)^2$ . Thus, by (32) we see that for all  $i \in [1 : n - n_0]$  we have the bound

$$\begin{aligned}
(36) \quad \mathbb{E}[d_{n,i}^2 \mid \mathcal{F}_{i-1}] &\geq \frac{1}{2}\rho^2 \mathbb{E}[(D'_i - D_i)^2 \mathbb{1}(G_i) \mid \mathcal{F}_{i-1}] \\
&= \frac{1}{2}\rho^2 \int_{s_\infty}^J \int_{s_\infty}^J (u' - u)^2 \psi(u') \psi(u) du \equiv \beta' > 0,
\end{aligned}$$

and  $\beta'$  is strictly positive because the first inequality of (23) implies that the density  $\psi$  cannot vanish identically on  $[s_\infty, J]$ .

Finally, using (36) in the range where it applies, we see that when we take total expectations and sum we obtain

$$\sum_{i=1}^n \mathbb{E}[d_{n,i}^2] \geq \sum_{i=1}^{n-n_0-1} \mathbb{E}[d_{n,i}^2] \geq \beta'(n - n_0 - 1).$$

When we take  $\beta = \beta'/2$  and  $N = 2(n_0 + 1)$ , the proof of the lemma is complete.  $\square$

From our decomposition (16) we have  $\mathcal{C}_n(\pi_n^*) = \mathcal{C}'_n(\pi_n^*) + \mathcal{C}''_n(\pi_n^*)$  where we also have  $\|\mathcal{C}''_n(\pi_n^*)\|_\infty \leq B(n_0 + 1)$  for a fixed constant  $B$ . Thus, by the Cauchy-Schwarz inequality we find

$$(37) \quad \text{Var}[\mathcal{C}_n(\pi_n^*)] = \text{Var}[\mathcal{C}'_n(\pi_n^*)] + O(\sqrt{\text{Var}[\mathcal{C}'_n(\pi_n^*)]}),$$

and from Lemma 9 we get that  $\text{Var}[\mathcal{C}'_n(\pi_n^*)] \sim \text{Var}[\mathcal{C}_n(\pi_n^*)]$  as  $n \rightarrow \infty$ . This completes the second — and last — step in the proof of the CLT for  $\mathcal{C}'_n(\pi_n^*)$ , i.e. we have

$$\frac{\mathcal{C}'_n(\pi_n^*) - \mathbb{E}[\mathcal{C}'_n(\pi_n^*)]}{\sqrt{\text{Var}[\mathcal{C}'_n(\pi_n^*)]}} \Longrightarrow N(0, 1) \quad \text{as } n \rightarrow \infty.$$

Finally, by the boundedness of  $\mathcal{C}''_n(\pi_n^*)$ , it is trivial that

$$\frac{\mathcal{C}''_n(\pi_n^*) - \mathbb{E}[\mathcal{C}''_n(\pi_n^*)]}{\sqrt{\text{Var}[\mathcal{C}'_n(\pi_n^*)]}} \Longrightarrow 0 \quad \text{as } n \rightarrow \infty,$$

so, by summing the last two relations and using (37) a second time, we complete the proof of Theorem 1 by applying Theorem 2 (or, more precisely, we apply the special case given by Corollary 3).

## 6. FURTHER CONNECTIONS AND OPEN PROBLEMS

In models where the optimality of a policy  $\pi_n^*$  is defined purely in terms of expected costs there are good reasons to be skeptical about the true economic value that is realized by the policy. A central limit theorem for the realized cost, like our Theorem 1 for the Bulinskaya inventory model, is one credible step toward the validation of the mean-optimizing technology. Nevertheless, there are other steps one could take.

In particular, a concentration inequality can be used to make a similar point. As we noted earlier in Remark 8, the martingale differences in our representation (29) have the uniform boundedness property (30). Consequently, by Hoeffding's inequality one finds for all  $\lambda \geq 0$  that

$$(38) \quad \mathbb{P}[|\mathcal{C}_n(\pi_n^*) - \mathbb{E}[\mathcal{C}_n(\pi_n^*)]| \geq \lambda] \leq 2 \exp\{-\lambda^2/(2nB)\}$$

where  $B = \sup\{\|d_{n,i}\|_\infty : 1 \leq i \leq n \text{ and } 1 \leq n < \infty\}$ . This bound gives one an independent argument that  $\mathcal{C}_n(\pi_n^*)$  is well represented by  $\mathbb{E}[\mathcal{C}_n(\pi_n^*)]$ . Specifically, it also tells us that deviations of order greater than  $O\{\mathbb{E}[\mathcal{C}_n(\pi_n^*)]^{1/2}\} = O(n^{1/2})$  have small, well-specified, probabilities.

The proofs of both Theorem 1 and the concentration inequality (38) make essential use of the assumption that the demand density  $\psi$  has bounded support. From a purely practical point of view, there is no problem with this assumption. Real-world demands are bounded. Nevertheless, in mathematical inventory theory one often allows for demand distributions that have unbounded support, so it is natural to ask if one has analogs of these results for such distributions.

The most direct approach to such an extension of Theorem 1 would be through the corresponding extension of the underlying CLT from Arlotto and Steele (2016) that we have used here. Such an extension is credible, but it is challenging. Alternatively, one might try to extend Theorem 1 by arguing from first principles, and here there is some encouragement from the unique features of the random variable  $\mathcal{C}_n(\pi_n^*)$ . In particular, one might hope to exploit the special structure of the martingale differences given by Proposition 7, but it is not easy to make more concrete suggestions.

On the other hand, since there are many flexible variations of Hoeffding's inequality, there are sunnier prospects for concentration inequalities for  $\mathcal{C}_n(\pi_n^*)$  when  $\phi$  has unbounded support. To be sure, such extensions are likely to lose the simplicity of (38), but they could be useful in concrete contexts.

Finally, there is a third approach to the economic viability of mean-optimal policies that can be built on the theory of stochastic orders. In the past, the theory of stochastic orders has been used to explain how properties of the transition kernels impact the the associated value functions (see, e.g. Müller, 1997; Smith and McCardle, 2002), but here we have something different in mind.

To be specific, we first recall that a random variable  $X$  is smaller than a random variable  $Y$  in the (usual) stochastic order if  $\mathbb{P}(X > t) \leq \mathbb{P}(Y > t)$  for all  $t \in \mathbb{R}$ , and, in this case, we write  $X \leq_{\text{st}} Y$ . Now, if we consider the policy  $\pi_n^*$  that is mean-optimal for the Bulinskaya model (or other stochastic optimization model), then it would be extraordinarily pleasing if one could prove that

$$(39) \quad \mathcal{C}_n(\pi_n^*) \leq_{\text{st}} \mathcal{C}_n(\pi) \quad \text{for all } \pi \in \mathcal{S}$$

where  $\mathcal{S}$  is a class of feasible policies.

One of the many consequences that would follow from (39) is that for each monotone increasing  $\phi$  one would have

$$\mathbb{E}[\phi(C_n(\pi_n^*))] \leq \mathbb{E}[\phi(C_n(\pi))] \quad \text{for all } \pi \in \mathcal{S}.$$

Since any utility function is monotone increasing, this would tell us that the mean-optimal policy  $\pi_n^*$  has *minimal expected utility cost* for any utility. To be sure, even this would not free one from the Saint Petersburg Paradox — essentially nothing does. Nevertheless, a bound like (39) would provide another important argument in the economic case for  $\pi_n^*$ .

For the moment, there is not quite enough evidence for (39) to put it forth as a conjecture. Still, this kind of question is understudied, and it provides a natural challenge. Moreover, there is considerable flexibility in how one might proceed. For example, to take the investigation a step further, one could consider any of the univariate orders in Shaked and Shanthikumar (2007) and ask again if one has the analog of (39) — either for the Bulinskaya model, or for other models that may be more tractable.

#### ACKNOWLEDGEMENT

This material is based upon work supported by the National Science Foundation under CAREER Award No. 1553274.

#### REFERENCES

- Arlotto, A., Gans, N. and Steele, J. M. (2014), ‘Markov decision problems where means bound variances’, *Oper. Res.* **62**(4), 864–875.
- Arlotto, A. and Steele, J. M. (2016), ‘A central limit theorem for temporally non-homogenous markov chains with applications to dynamic programming’, *Math. Oper. Res.* **41**(4), 1448–1468.
- Bulinskaya, E. V. (1964), ‘Some results concerning optimum inventory policies’, *Theory of Probability & Its Applications* **9**(3), 389–403.
- Colangelo, A., Scarsini, M. and Shaked, M. (2005), ‘Some notions of multivariate positive dependence’, *Insurance Math. Econom.* **37**(1), 13–26.
- Dobrushin, R. L. (1956), ‘Central limit theorem for nonstationary Markov chains. I, II’, *Theory of Probability & Its Applications* **1**(1), 65–80; *ibid.* **1**(4), 329–383.
- Hartman, B. C., Dror, M. and Shaked, M. (2000), ‘Cores of inventory centralization games’, *Games Econom. Behav.* **31**(1), 26–49.
- Li, X. and Shaked, M. (2007), Stochastic orders based on total time on test transform, in ‘Advances in statistical modeling and inference’, Vol. 3 of *Ser. Biostat.*, World Sci. Publ., Hackensack, NJ, pp. 131–143.
- Marshall, A. W. and Shaked, M. (1983), ‘New better than used processes’, *Adv. in Appl. Probab.* **15**(3), 601–615.
- Müller, A. (1997), ‘How does the value function of a Markov decision process depend on the transition probabilities?’, *Math. Oper. Res.* **22**(4), 872–885.
- Müller, A., Scarsini, M. and Shaked, M. (2002), ‘The newsvendor game has a nonempty core’, *Games Econom. Behav.* **38**(1), 118–126.
- Sethuraman, S. and Varadhan, S. R. S. (2005), ‘A martingale proof of Dobrushin’s theorem for non-homogeneous Markov chains’, *Electron. J. Probab.* **10**, no. 36, 1221–1235 (electronic).
- Shaked, M. and Shanthikumar, J. G. (1994), *Stochastic orders and their applications*, Probability and Mathematical Statistics, Academic Press, Inc., Boston, MA.
- Shaked, M. and Shanthikumar, J. G. (2007), *Stochastic orders*, Springer Series in Statistics, Springer, New York.
- Shaked, M., Sordo, M. A. and Suárez-Llorens, A. (2012), ‘Global dependence stochastic orders’, *Methodol. Comput. Appl. Probab.* **14**(3), 617–648.
- Smith, J. E. and McCardle, K. F. (2002), ‘Structural properties of stochastic dynamic programs’, *Oper. Res.* **50**(5), 796–809.